

DATA CATALOG SIDE MEETING

ŠOLTA, 22 AUGUST 2011

ESSI-Lab, Italian National Research Council

Enrico Boldrini, Lorenzo Bigagli, Stefano Nativi

enrico.boldrini@cnr.it

lorenzo.bigagli@cnr.it

stefano.nativi@cnr.it



Air Quality
Community of Practice

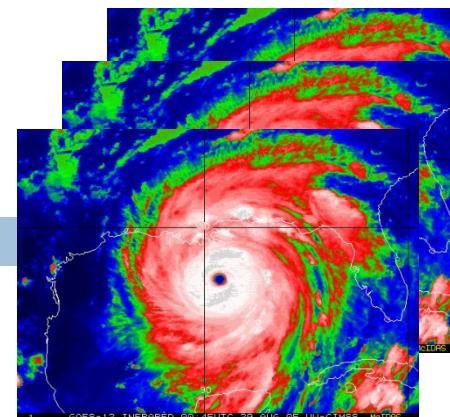
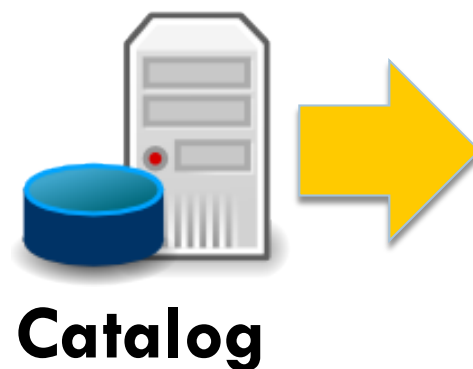


Presentation overview

- Catalog introduction
 - ▣ Catalog architectures (centralized, distributed, broker)
- A broker catalog: GI-cat
 - ▣ Components & features
- Air Quality Community catalog Interoperability tests
 - ▣ Overview & demo
- Issues & discussion

Data Catalog overview

3 / 17



Metadata for presentation/use

Metadata elements

- **Metadata for discovery**
used to formulate queries,
generic enough to embrace
different disciplines/domains
- **Metadata for presentation/
use**
describes the dataset
belonging to a given domain
at the maximum detail level

CSW ISO mandates ≈ 30 queryable
elements, amongst them:

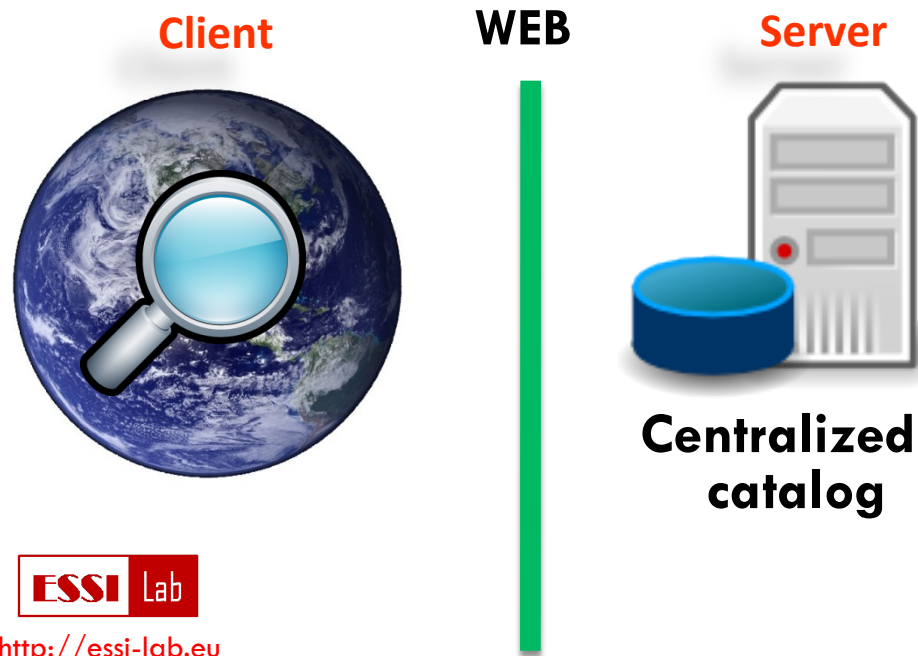
- TopicCategory
- FileIdentifier
- RevisionDate
- ...

Name	Value
...	...
Distributor info	
Contact info	
OrganizationName	DataFed
Address	
Delivery point	1 Brookings Dr
City	Saint Louis
Administrative area	MO
Postal code	63130
Country	USA
Online resource	
Linkage	http://datafed.net
Role	distributor
Transfer options	
Digital transfer options	
...	...

Catalog architectures

5 / 17

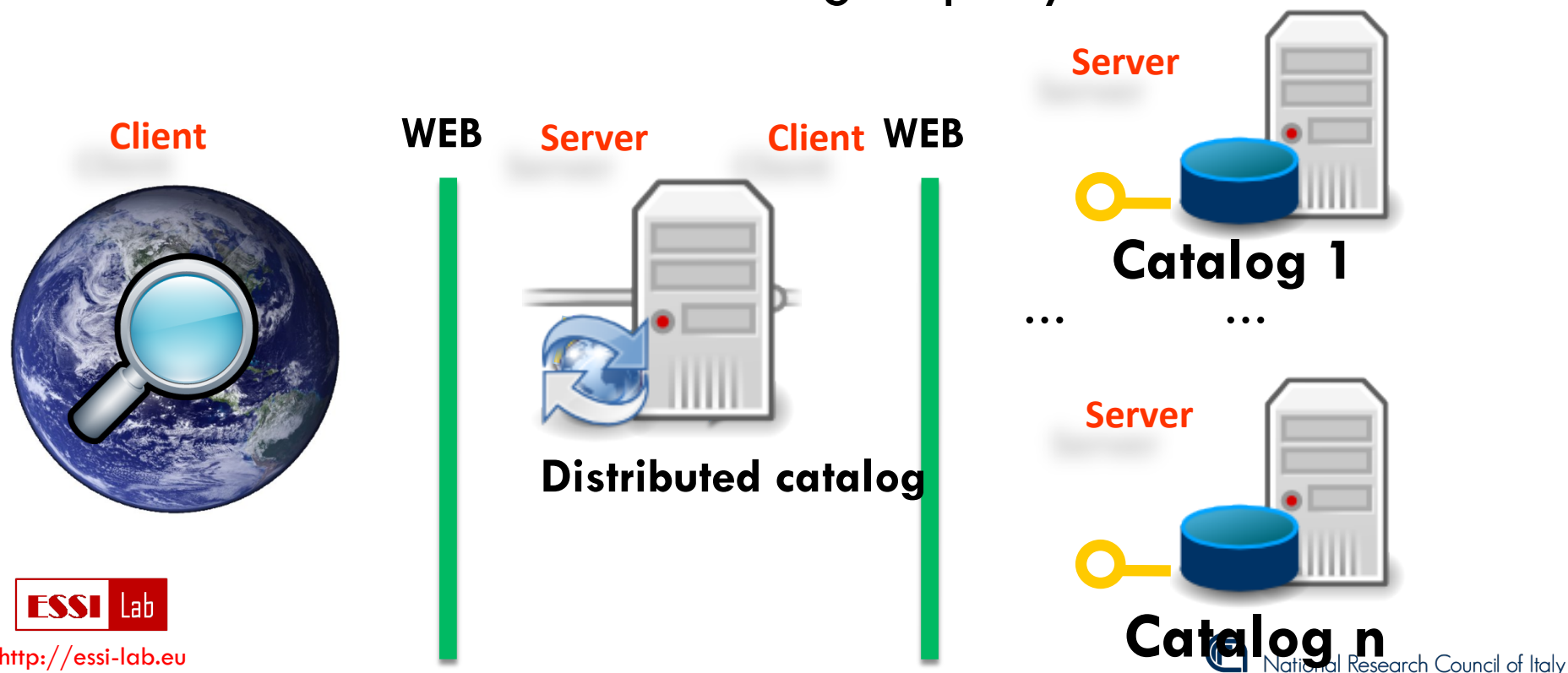
- **Centralized catalog:** all the available information is stored on a local offline database



Catalog architectures

6 / 17

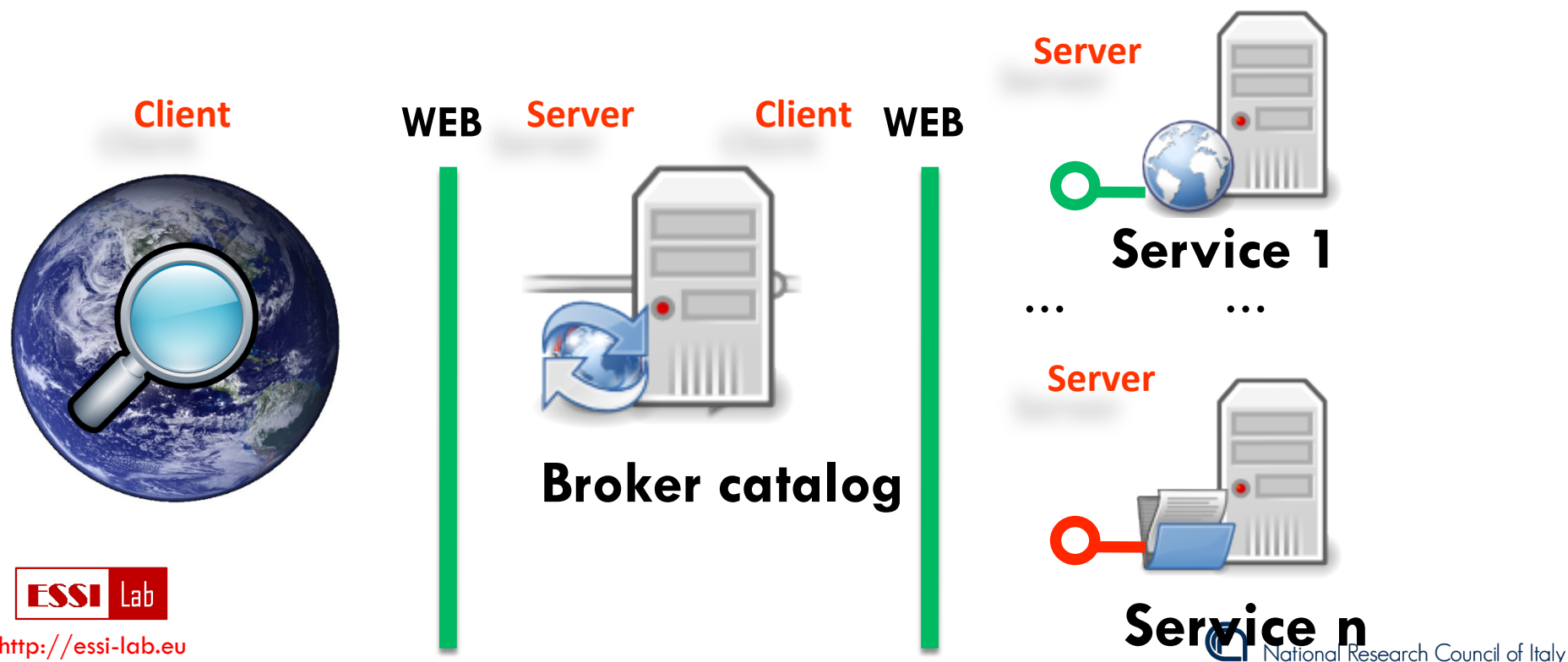
- **Distributed federated catalog:** transparently to the catalog user, multiple catalogs are queried in order to obtain results from a single query



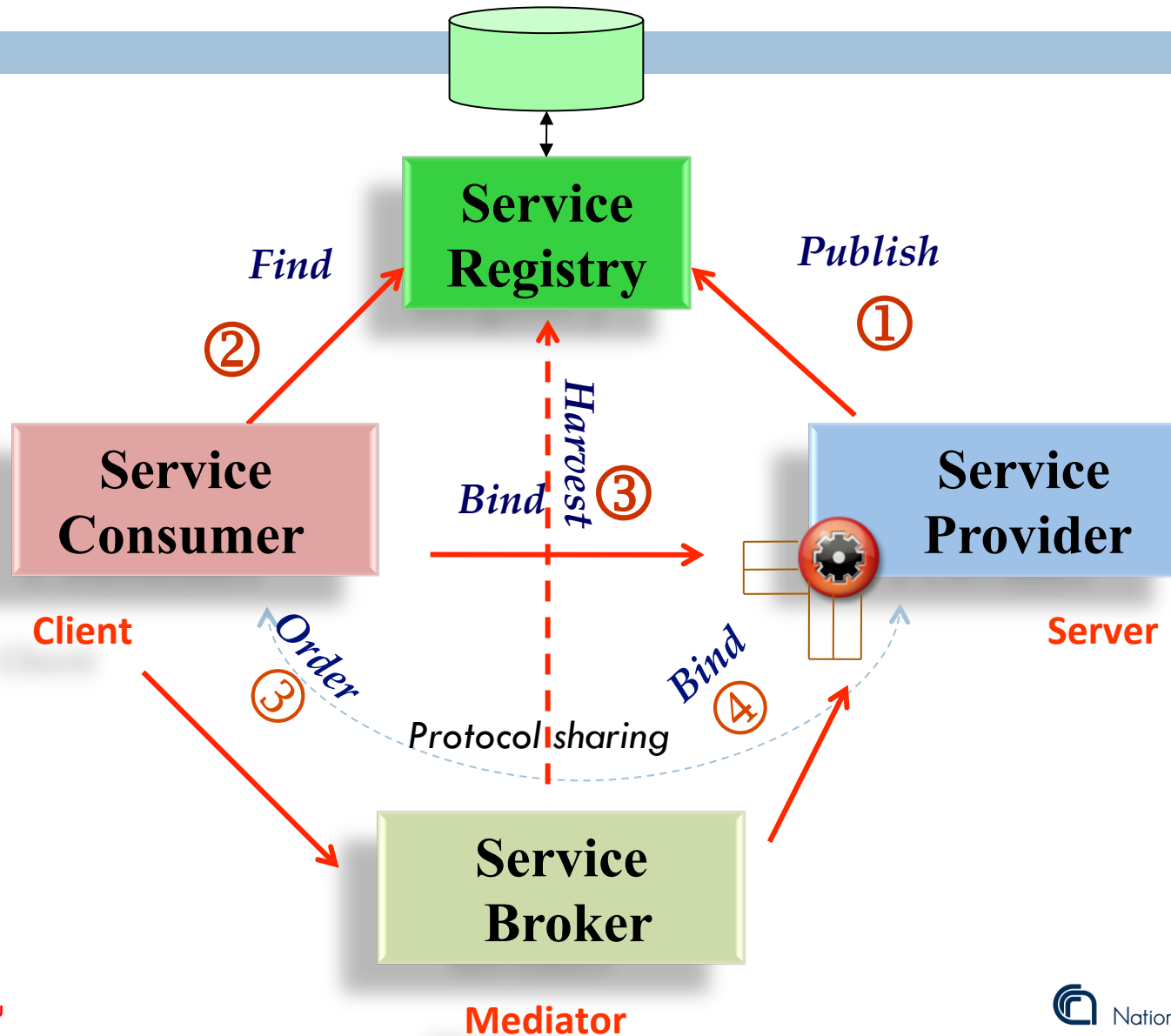
Catalog architectures

7 / 17

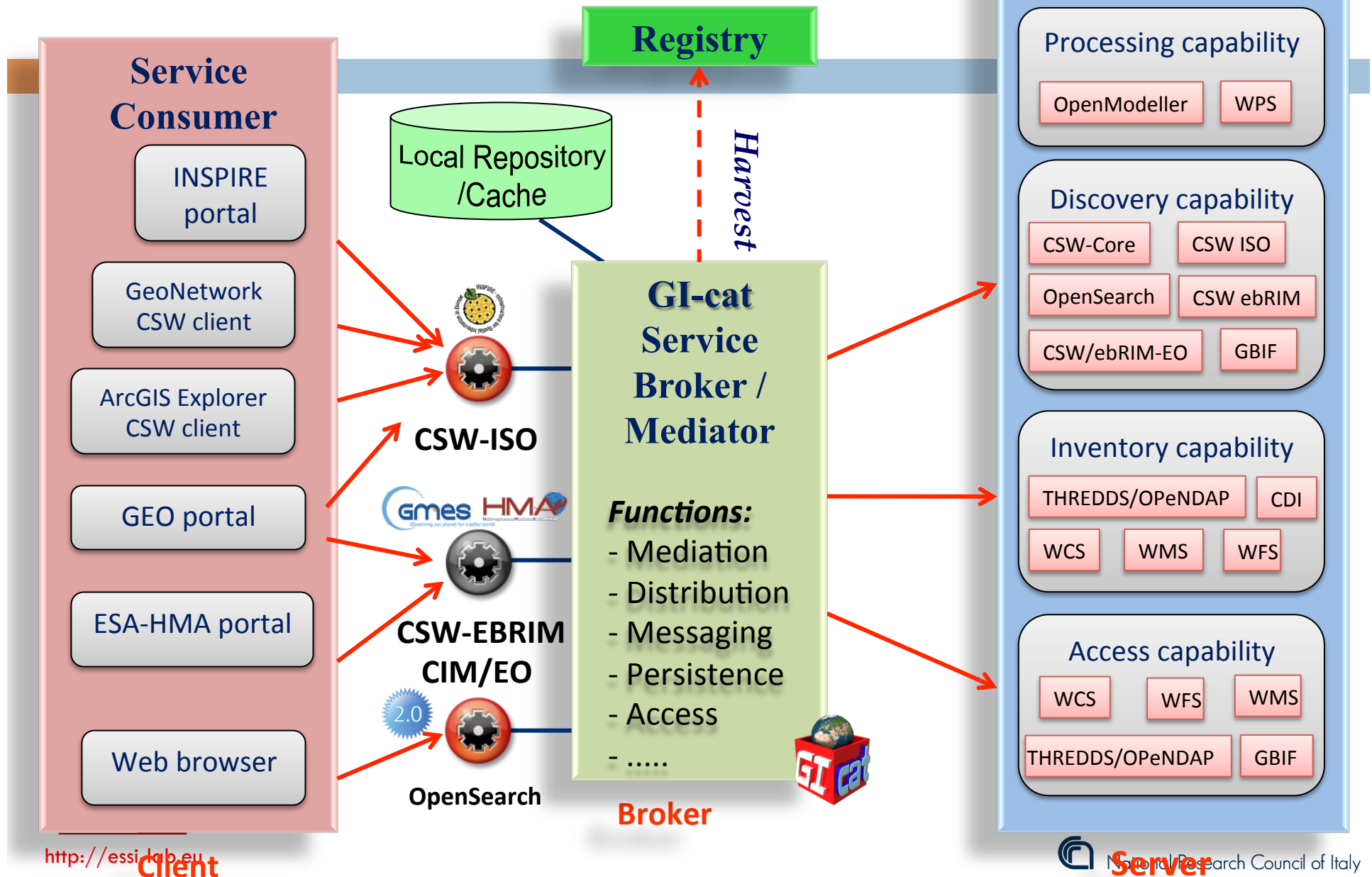
- **Broker catalog:** transparently to the catalog user, **heterogeneous services** are queried in order to obtain results from a single query

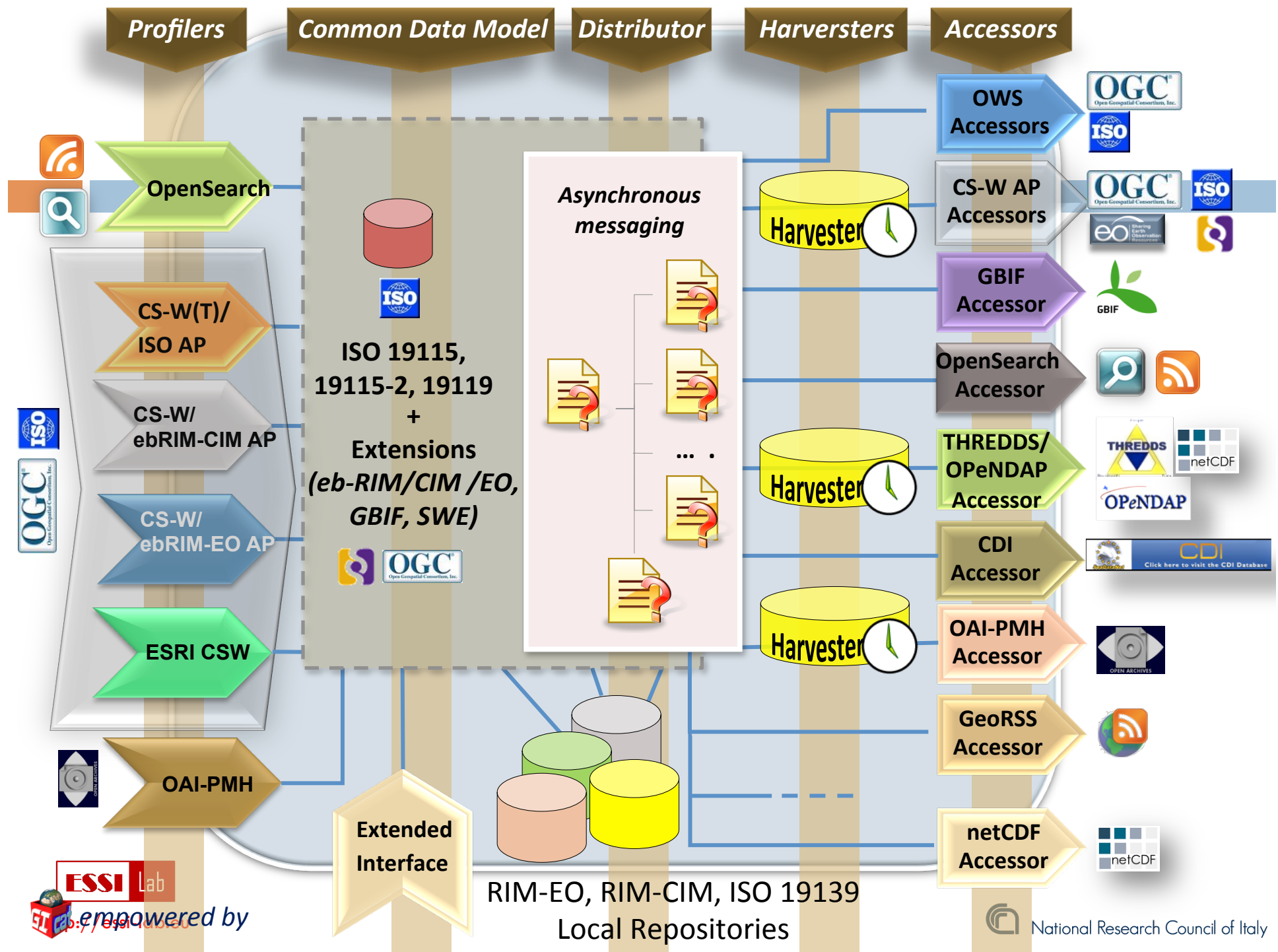


A SOA Broker/Mediator component



GI-cat Broker/Mediation component





Available Accessors

Discovery/Inventory/Listing services



WCS



WMS



WPS



WFS



SOS



CSW

CORE



OGC WCS 1.0, 1.1, 1.1.2

OGC WMS 1.3.0, 1.1.1

OGC WFS 1.0.0

OGC WPS 1.0.0

OGC SOS 1.0

OGC CSW 2.0.2 Core, AP ISO 1.0, ebRIM/CIM, ebRIM/EO

ESRI ArcGIS Geoportal (version 10) catalog service

THREDDS 1.0.1, 1.0.2

THREDDS-NCISO 1.0.1, 1.0.2

CDI 1.04, 1.3, 1.4 1.6

GI-cat 6.x, 7.x

GBIF

Available Accessors

Discovery/Inventory/Listing services



GeoNetwork (versions 2.2.0 and 2.4.1) catalog service



Deegree (version 2.2) catalog service



OpenSearch 1.1 accessor



OAI-PMH 2.0 (support to ISO19139 and dublin core formats)

Available Accessors

Data Handlings



NetCDF-CF 1.4



NCML-CF



NCML-OD



ISO19115-2



GeoRSS 2.0



GDACS



DIF



SITAD (Sistema Informativo Territoriale Ambientale Diffuso)



Generic File system



Web Accessible Folders

Available Profilers

Discovery/inventory services



OGC CSW



CSW 2.0.2 AP ISO 1.0



CSW 2.0.2 ebRIM EO



CSW 2.0.2 ebRIM CIM



ESRI GEOPORTAL 10



OAI-PMH 2.0



OpenSearch



OpenSearch 1.1



OpenSearch GENESI DR



GI-cat extended interface



<http://essi-lab.eu>



National Research Council of Italy

Interoperability through the use of standards (De iure/de facto)

□ Data model standards

- ISO 19115/ISO 19139



- Dublin core



□ Standard services

- FTP, HTTP, WAF, ...

- WCS, WMS, WFS, ...

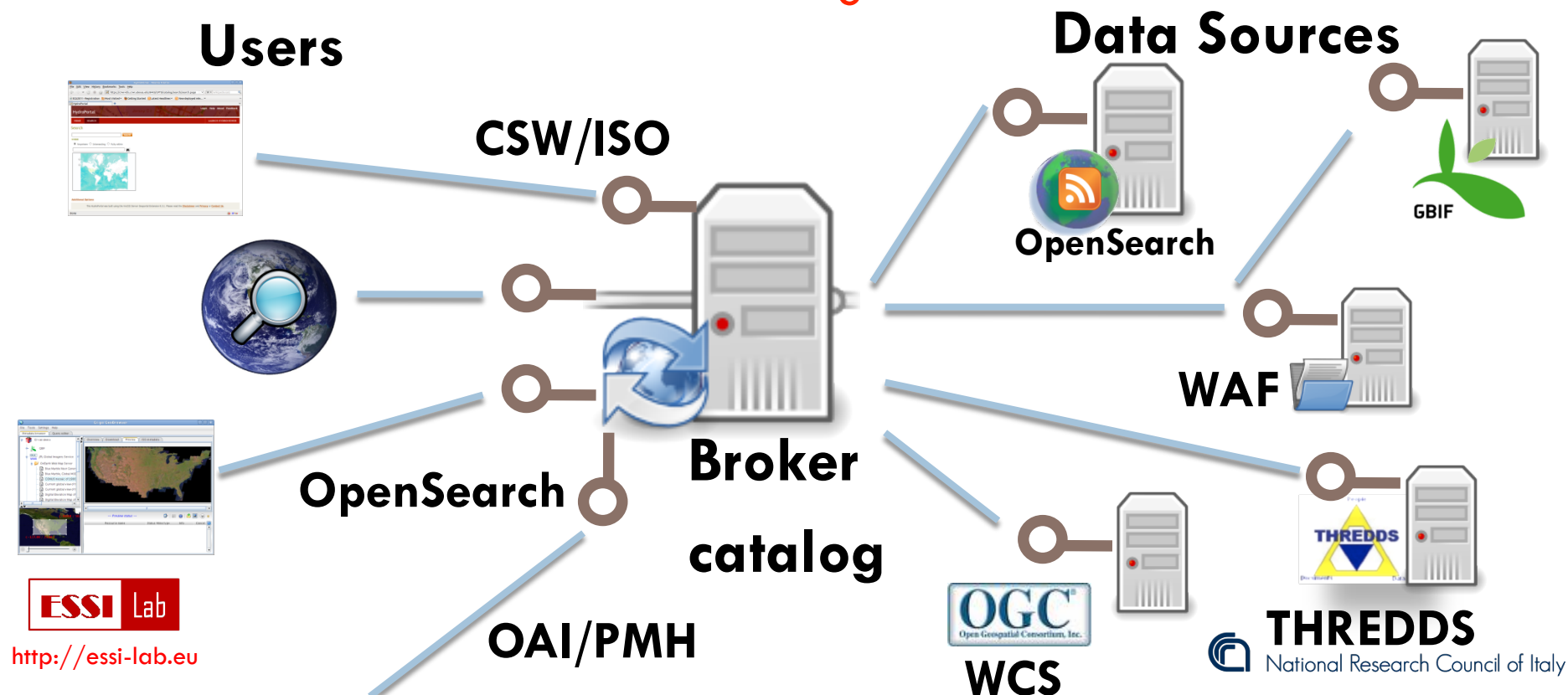


- THREDDS, CSW/ISO, OAI-PMH, OpenSearch, ...



Broker catalog use cases

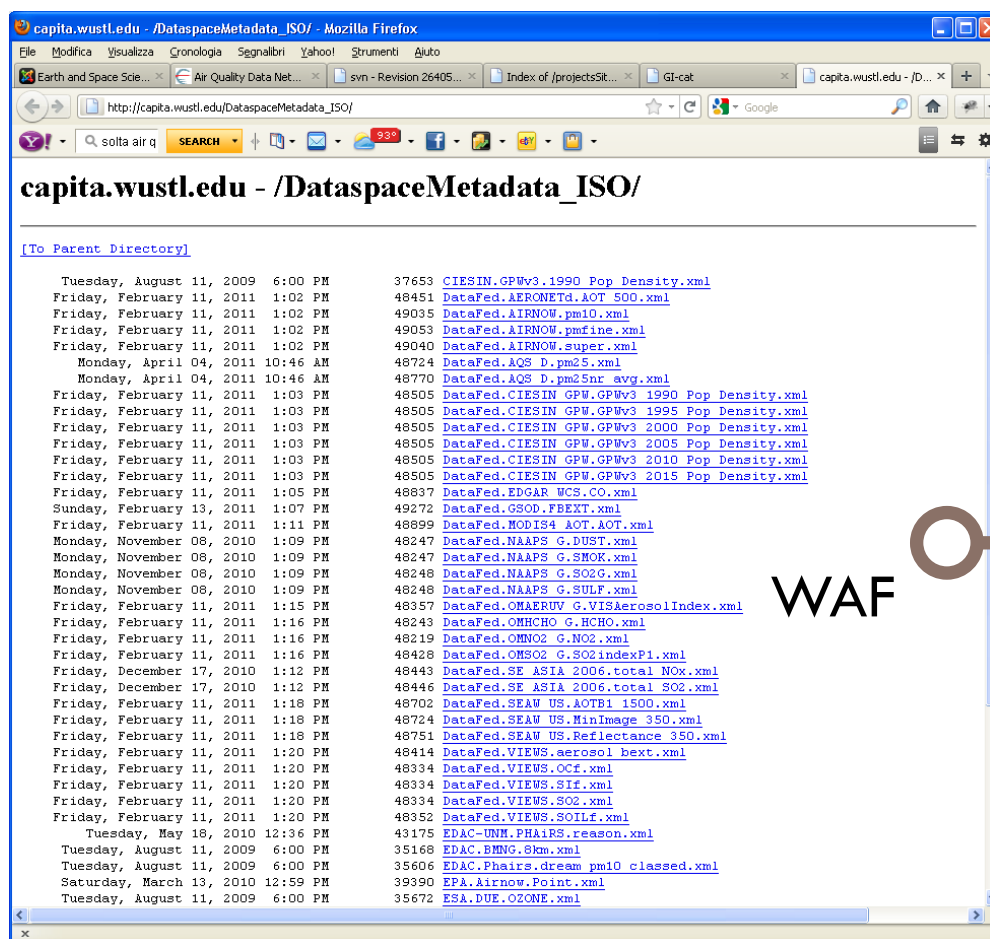
- Adding **new heterogeneous nodes** to the existing capability
→ **new data is made available**
- Adding **new standard catalog interfaces**
→ **the data is made available to existing and even future clients**



AQ Interoperability tests

- ❑ GI-cat federating the Air Quality Community catalog
 - ▣ Browsing and querying of the published data has been enabled by the use of standard catalog interfaces (demo with different catalog clients)

Web Accessible Folder - WAF



uFind web tool

WAF

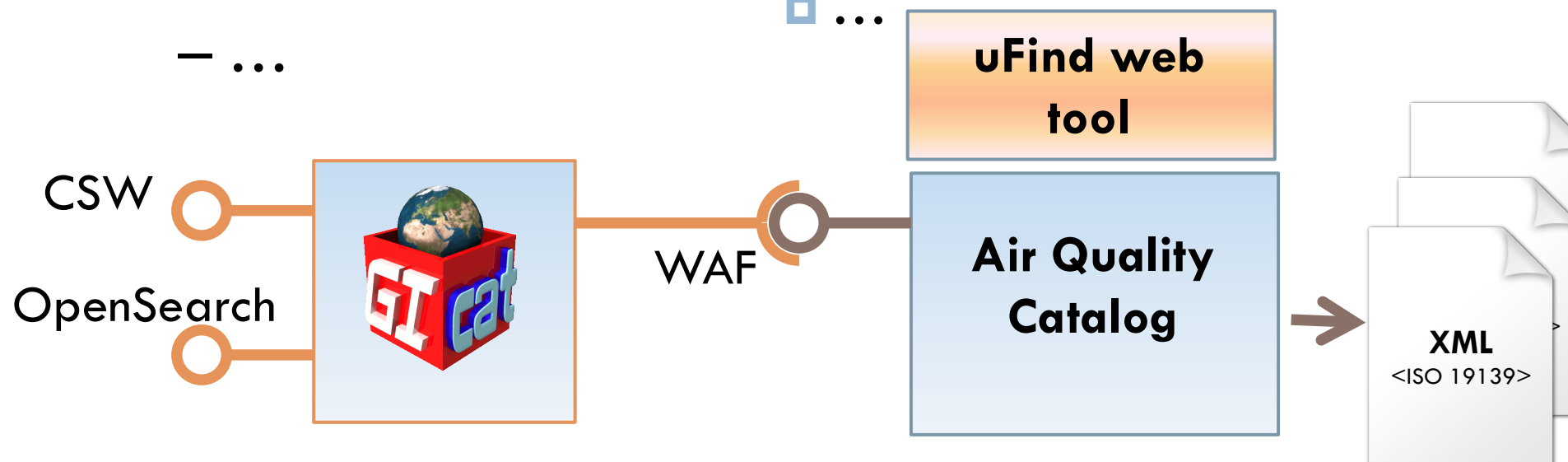
Air Quality Catalog

XML
<ISO 19139>

- GI-cat Output
 - CSW ISO 2.0.2
 - OpenSearch
 - OAI-PMH
 - ...

- GI-cat Input:

- Web Accessible Folders -WAF
- Local Folders
- WCS, WMS, W*S
- ...

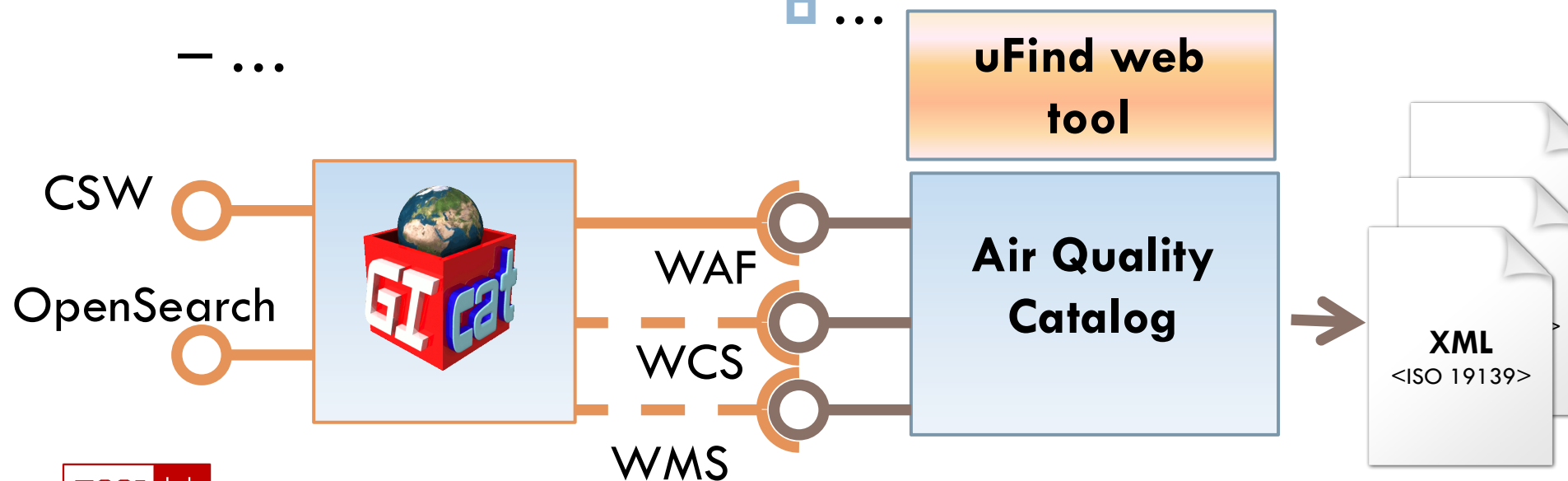


- **GI-cat Output**

- CSW ISO 2.0.2
- OpenSearch
- OAI-PMH
- ...

- **GI-cat Input:**

- Web Accessible Folders -WAF
- Local Folders
- WCS, WMS, W*S
- ...



Interoperability with different clients

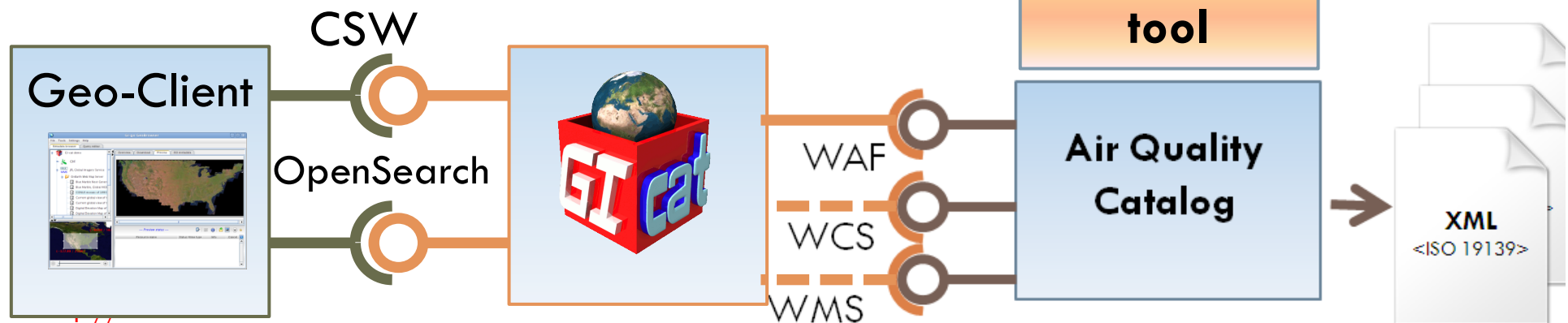
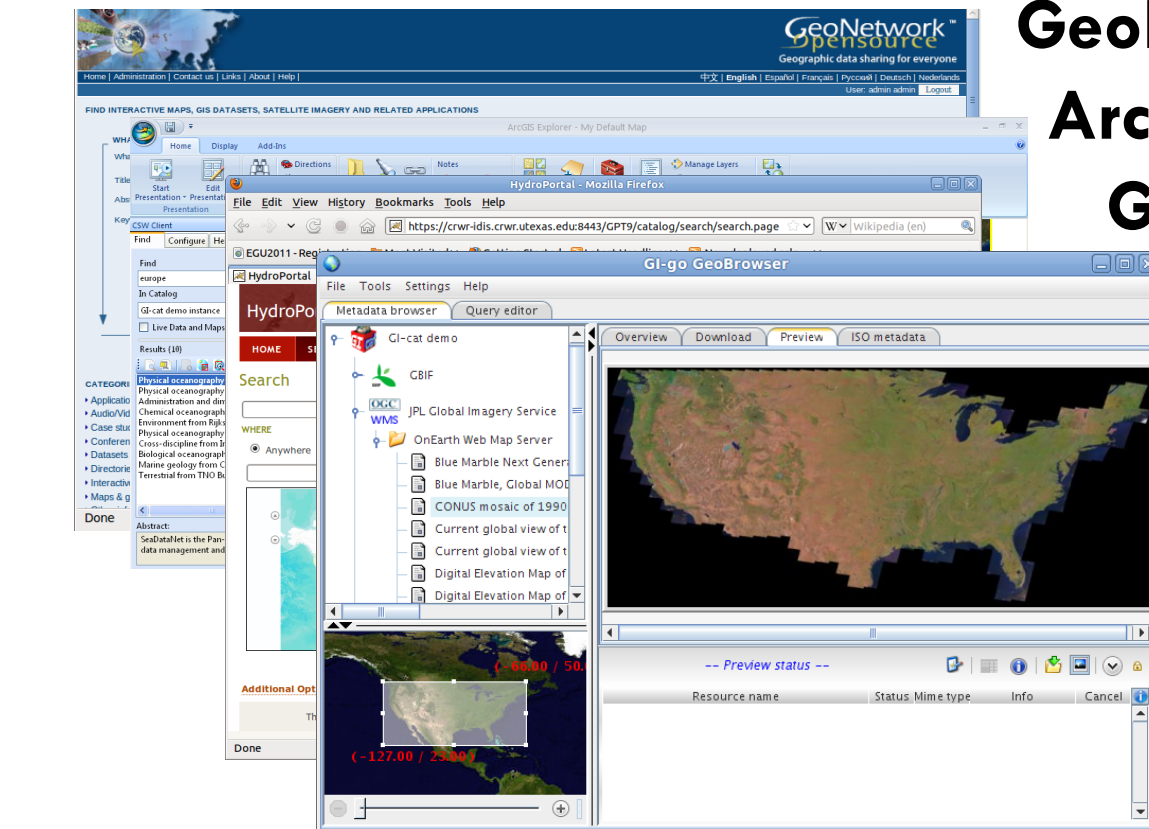
GeoNetwork CSW client

ArcGIS Explorer CSW client

GeoPortal

GI-go

Geobrowser / GI-portal



CSW Browsing using GI-go client

The screenshot displays the GI-go GeoBrowser application. The interface is divided into several sections:

- Metadata browser (left):** A tree view showing a list of datasets under the 'WebAccessibleFolder'. The selected dataset is 'AIRNOW.pm10'.
- Map (bottom left):** A world map showing the location of the selected dataset, with a red box indicating the bounding box coordinates (30.00 / 24.00).
- Detailed view (right):** A table showing the metadata for the selected dataset, organized into sections like 'Descriptive keywords', 'Resource constraints', 'Spatial representation type', 'Extent', 'Temporal element', and 'Service identification'.

Name	Value
Type	theme
Descriptive keywords	
Keyword	Units:ug/cm^3
Type	theme
Descriptive keywords	
Keyword	AQCommunityCatalog
Type	theme
Descriptive keywords	
Keyword	Project:ACC
Type	theme
Resource constraints	
Constraints	
Use limitation	no conditions apply
Resource constraints	
Legal constraints	
Access constraints	otherRestrictions
Other constraints	no limitations
Language	English
Topic category	climatologyMeteorologyAtmosphere
Spatial representation type	
Code	grid
Spatial resolution	
Denominator	555
Character set	utf8
Extent	
Geographic element	
Geographic bounding box	
West	-130
East	-65
South	24
North	52
Temporal element	
Begin	2002-07-01
End	2011-02-12
Service identification	
Title	AIRNOW DataFed Web Map Server 1.1.1
Date	2002-07-01
Date type	creation
Abstract	AIRNOW DataFed Web Map Server 1.1.1
Type	urn:ogc:serviceType:WebMapService:1.1
Coupling type	tight
Operation	

http://capita.wustl.edu/DataspaceMetadata_ISO/DataFed.AIRNOW.pm10.xml

Query using GI-portal Opensearch – online demo at: <http://ec2-174-129-9-172.compute-1.amazonaws.com/gi-cat-9>

The screenshot shows the GI-portal Opensearch web application. The main map displays North America with a yellow selection box over the United States. The search results table shows 7 results for the query 'air'. The query constraints panel on the right allows for keyword, location, and selected area selection.

Spatial constraint selection and results localization

QUERY CONSTRAINTS SELECTION

- Keyword
air

- Location
Enter a location name (case is ignored), e.g.: europe,italy,rome,etc...

- Selected area
61.754
-147.689
-51.713
12.535

- Time
From:
To:

- Results per page
10

Start search

Search results: 7 - Elapsed time: 1 seconds

Metadata	Title	BBOX	Layer	Select/Unselect all
View metadata	AIRNOW.pm10	Zoom		<input type="checkbox"/>
View metadata	AIRNOW.pmfine	Zoom		<input type="checkbox"/>
View metadata	AIRNOW.super	Zoom		<input type="checkbox"/>
View metadata	AQS_D.pm25nr_avg	Zoom		<input type="checkbox"/>
View metadata	AQS_D.pm25	Zoom		<input type="checkbox"/>

Find: 9.0.1 Previous Next Highlight all Match case

Next steps (beyond discovery...)

□ Data Access

- ▣ OGC services (WCS, WMS, WFS, SOS)
- ▣ metadata conventions (CI_OnlineResource) for OWS data access
- ▣ brokered-access (EuroGEOSS access broker / GI-axe)

GI-cat homepage

<http://essi-lab.eu/gi-cat>

Download an **autoinstaller package** for an easy default installation;

- Windows
- Linux
- Mac OS X

Download **Java .war package** for custom deploy scenario.

User mailing list

<http://essi-lab.eu/gi-devel>



<http://essi-lab.eu>



Issues/discussion

- Discovery Metadata in AQ (are queryable properties of CSW AP ISO sufficient? / CSW queryables review - ISO properties for discovery)
- AQ data granularity level
- How can the AQ Community Catalog and Gl-cat complement each other effectively?
 - Interoperability through standard interfaces and data models (WAF, ISO 19139, ...)
 - uFind functionalities and features, investigate the possibility to use it as a Gl-cat client...



Thanks for your attention!